

## IDENTITY CONSTRUCTION THROUGH CODE SWITCHING: A SOCIO-LINGUISTIC ANALYSIS OF STUDENTS' LANGUAGE USE

Aysha Parveen Noor<sup>\*1</sup>, Nazia Anwar<sup>2</sup>, Sara Ilyas<sup>3</sup>

<sup>\*1,3</sup>MPhil Scholar, English, University of Gujrat

<sup>2</sup>Assistant Professor, English, University of Gujrat

<sup>1</sup>25011702-016@uog.edu.pk, <sup>2</sup>nazia.anwar@uog.edu.pk, <sup>3</sup>ilyassara594@gmail.com

DOI: <https://doi.org/10.5281/zenodo.19882738>

### Keywords

### Article History

Received: 01 March 2026

Accepted: 10 April 2026

Published: 29 April 2026

Copyright @Author

Corresponding Author: \*

Aysha Parveen Noor

### Abstract

This study aims to investigate the matrix language in English-Urdu code-switching among the Pakistani students and how they perform their linguistic identity through that. This present study further extends to explore the correlation between education, matrix language and identity construction. The study has adopted the qualitative approach, examining the intermediates, undergraduates and graduates in Pakistan through Myers-Scotton Matrix Language Frame Model. The findings have revealed that Urdu functions as the matrix language while English serves the content morphemes among the intermediates, reflecting locally grounded identity. In contrast, among the under-graduates English inconsistently emerges as the matrix language, indicating the transitional linguistic identity, while among the graduates, English consistently appears as the matrix language, corresponding to the stable academically grounded and English-dominant linguistic identity. Moreover, education appeared to be an effective social factor, shifting the matrix language choice among the students, which indicated a correlation between identity construction through matrix language and education.

**Keywords:** Matrix language, code-switching, education, Pakistan, embedded language, identity construction

### 1. INTRODUCTION

Bilingualism is the competence of more than one language. Beardmore (1982) defines bilingualism as the “ability of a person to function, without any traces of the language in his use of the other” and he perform perfect in both languages in all the domains of activity, but Huttner (1997) did not support this idealistic definition of bilingualism and said it overlooks those bilinguals who are more proficient in one language than the other. Grosjean (2014) provides the perfect definition of bilingualism, and called it the mixture of two languages, which allows for adjustable communication in various communicative

atmospheres. On the other hand, Wei, L (2000) argued that while describing and defining bilingualism, the most challenging step is to explain the extent to which bilingualism is practiced by bilingual speakers. Code switching is the process of shifting between two or more languages. It frequently happens among bilingual speakers. Hymes (1974) added that code-switching is the name of frequent alternation between two languages, linguistic forms and the style of speaking. Building upon this, Heller (1998) argued that code switching is the use of more than one language within a single communicative episode. Gumperz (1977) added that code-switching is a

part of minority groups' informal speeches in society.

Moreover, Gumperz (1982) defines code-switching as not a deficit approach to linguistics but a discourse strategy used by most speakers to express social meaning in which they combine two languages that belong to different social groups. Moreover, Crystal (1992, p 60) and Chaer (1995, p 141) stated that during the conversation speakers tend to change between the two languages, which is called code-switching. While Graddol, Leith and Swann (1996) called code-switching an ambiguous choice. Poplack (1980) stated that code-switching can be divided into tag switching, intra-sentential and inter-sentential code-switching and Hammink (2000) added that if there is no morphological adaptation, then intra-sentential code-switching occurs at the clause or phrase level. In the study of code-switching, the most influential approach is Myers-Scotton Matrix Language Frame Model (1993), which emphasizes that code-switching is not random but structurally constrained. In intra-sentential code-switching, one language acts as the matrix language, which controls and dominates the grammatical structure. Matrix language provides the system morphemes and decides the morpheme order of the clause. Another language acts as the embedded language which provides the content morphemes that have to fit in the morpheme order of the matrix language.

Apart from the structural dimension, code-switching plays a significant role in the construction and expression of identity especially in multilingual societies where the speakers systematically alternate between languages to position themselves socially and negotiate identities that may be local, global or hybrid in nature. Hence, language choice is not merely a phenomenon of communication but a social practice of framing and performing identities. In Pakistan, Urdu and English exist simultaneously, where Urdu serves as the national language while English serves as the official and international language. Urdu is mostly used for communication and English carries more prestige and is used in education, administration and other formal settings.

A lot of empirical studies have focused on the attitude and pragmatic motivations regarding code-switching while little attention has been paid to the identity construction through the matrix language choice across the three Pakistani educational levels: intermediates, undergraduates and graduates and the interrelation between identity construction through matrix language choice and education. So, the present study is focused on exploring this phenomenon. The study employed the Matrix Language Frame Model's two principles: morpheme order and system morpheme principle for the analysis of the English-Urdu code-switching.

### 1.1: Research Objectives

- To explore the variations in matrix language choice in English-Urdu code switching and its role in constructing the identities among the Pakistani students.
- To examine the relationship between the identity construction, matrix language choice and education.
- To analyze how far does the Myers-Scotton Matrix Language Frame Model adequately explain the English-Urdu code-switching.

### 1.2: Research Questions

1. What matrix language emerges among the Pakistani bilinguals and how does their linguistic choices contribute to identity construction?
2. To what extent does the Myers-Scotton Matrix Language Frame Model adequately explain the English-Urdu code-switching?

## 2. Literature Review

The study aims to examine the relationship identity construction through code-switching and the influence of education on the matrix language choice. The topic of English-Urdu code switching has gained significant attentions in recent years within the field of sociolinguistics. Hence, this literature review aims to critically examine these contributions to identify the research gap.

### 2.1 Identity Construction Through Language

Language is not merely a tool for communication but it's a medium through which individuals

actively construct, negotiate and express identities. Identities are viewed as fluid and dynamic in sociolinguistic and through language individual consistently shape and reshape their identities to show particular group alignment. Bucholtz and Hall (2005) argued that identities are produced through the interaction and are linked to language use. This framework indicates that language choice carries social meanings and allows the speaker to align themselves to the particular group. Building upon this, De Fina (2007) argues that discourse and interaction are the main elements for constructing identities. Speakers consistently use linguistic resources to present themselves in a particular way. Moreover, the correlation between language and identity becomes more significant in multilingual societies where speakers strategically alternate between languages to express different forms of identities. Heller (2007) suggested that linguistic choice in bilingual settings is inextricably linked with social practice and the source of identity construction and social positioning.

### 2.2 Social Factors and Code-Switching

Social factors such as education, gender, motivation, etc. play a huge role in the code-switching. People switch code due to the influence of factors such as “context and situation”, “participant”, “topic” and “important characteristic of participant” (Pride, 1979, p 131). Similarly, Cormack (1979, p 18) code switching occurs due to the influence of “ethnic linguistic necessity”, “linguistic ability”, “the user’s goal” and “personality, circumstances and society”. Additionally, Holmes (2001, p 12) particular social factors such as social context in which the communication occurs, the specific function language performs in that social context and the topic of the communication have a greater influence on the linguistic choices. Hadie et al (2016) found the impact of “show identity” on the Malaysian-English code-switching. Similarly, Khan et al (2021) examined the Pashto-English code-switching and social factors and found the influence of education as the major factor in perpetuating the English usage. It indicates that education can create more bilinguals in Pakistan leading to the code-switching. Moreover,

Usmanova (2024) emphasized that social factors such as identity, social relationship, topic, context, emotional expression, bilingual competence, social or cultural norms and pragmatic efficiency are the major drivers of code-switching. Additionally, Prosper (2025) founded that contextual factors have an influence on code-switching and they should be taken into consideration in educational settings to advance the communicative competence of the learners.

### 2.3 English-Urdu Code Switching in Pakistan

Pakistan is a land that is rich in linguistic varieties. It has 70 to 80 regional languages, one national language, Urdu and one official and international language, English. The most dominant languages in Pakistan are Urdu and English because these are administrative, educational and media languages, which leads to the emergence of code-switching between English and Urdu. Rahman (1990) added that Pakistanis borrow words from the Urdu language speaking the English due to the influence of religion, Pakistani language and culture, history, Arabic and Pashto. Building upon this, Zainab, Ahmed and Kashif (2024) highlighted that code-switching in Pakistan occurs to fill the lexical gap. Asghar, Shehzad and Hanif (2023) debated that Britishers came to the sub-continent not only with the motive of prevailing English language but to replace the whole lifestyle of the people with the Britishers’ life style, calling it more civilized and modern. So, English in Pakistan is not just treated as a language, but a maker of social identity, class and intelligence, as Iqra, Mahmood, Firdaus and Saqlain (2025) stated that the reasons behind code-switching in Pakistani writers’ work are more than just linguistic requirements, in fact, it indicates the influence of globalization and the prestige associated with English.

In Pakistan, people have different attitudes towards Urdu-English code-switching Tahir, Fatima and Abuzar (2016) found that teachers in Pakistan prefer code-switching when the students find it difficult to understand the concept in a single language, and students also agreed that they understand the concepts better when the teacher switches between two codes. Similarly, Tabbasum (2023) added that Pakistani students have a very

positive attitude towards Urdu-English code-switching because it serves the preservation of cultural identity and helps them to understand the concepts more easily, making the learning easier. On the other hand, Fatima, Mustafa and Amin (2025) found that some students have the same positive attitudes towards code-switching while some students showed negative attitudes towards Urdu-English code-switching because reliance on English-Urdu code-switching hinders their linguistic development.

### 2.4 Research Gap

There is significant research on the English-Urdu code-switching. But little attention has been paid to the matrix language choice across the three different Pakistani educational levels (intermediate, undergraduates and graduates) and how the students negotiate their identity through code-switching. Moreover, fewer studies have explored the relationship between education, matrix language choice and identity construction. To combat this gap, this study aims to examine the matrix language variation at different educational levels and the relationship between education, matrix language choice and negotiation of different kinds of identities.

## 3. Methodology

### 3.1 Research Design

The researcher adopted the qualitative approach and built the study on interpretivism for analyzing the matrix language in the English-Urdu code-switching across different educational levels and negotiation of identity. It further extends to explain the relationship between education, matrix language and identity construction.

### 3.2 Variables of the Study

The present study includes three main variables: independent variable, mediating variable and dependent variable. These variables are operationalized to examine the relationship between education, matrix language and identity construction. Firstly, education is treated as the independent variable and used to examine how this influences the language use among the speakers. Secondly, the mediating variable is matrix

language choice, which explains how the speakers alternate between Urdu and English to construct identity. Lastly, the dependent variable is linguistic identity construction, which is analyzed through the linguistic preferences.

### 3.3 Data Collection

The researcher has selected 18 students, 6 from each educational level, from the intermediate, undergraduates and graduates as a sample to analyze the matrix language in code-switching. Purposively 6 students from each educational level were selected who knew both English and Urdu. To ensure the students' sufficient capacity in both languages, the researcher took the guidance of their instructors. Students were selected from ESL classrooms. From the intermediate level students of age group 17-19, from the undergraduate level students of age group 20-23 and from the graduate level students of age group 24-26 were selected to ensure the age gap of students of different levels. After sampling, the participants were asked to speak on any topic which suits to their interests. The researcher used notes and an audio-recorder to record the data. Then, the collected data was transcribed into written form. Moreover, the researcher conducted a semi-structured interview with them regarding their choice of certain matrix language to understand the impact of the social factor: education. After reading and re-reading the transcribed data, the researcher applied the Matrix Language Frame Model on the collected data to identify the matrix language in the student's English-Urdu code switching and how they negotiate identity through the code-switching. The study employed the two main principles of the MLF model: the morpheme order principle and the system morpheme principle for the analysis of the data. Only selected excerpts from the participants' code-switched clauses were examined in detail as the illustrative example to demonstrate the structural and functional features of code-switching. Rather than analyzing each and every example, most representative instances of broader patterns observed across the dataset were chosen, and recurring discourse and morphosyntactic patterns were identified.

### 3.4 Ethical Considerations

The study followed all the ethical principles. All the students who had participated in the study were clearly informed about the objective, nature and procedure of the study. There were no enforced interview and students' willingness to participate was ensured. Moreover, in case the students did not want to continue the participation and felt uncomfortable, they were given free hand to withdraw from the investigation at any point without any hesitation. The researcher ensured the confidentiality and anonymity of their names, institutions and recordings. The data was collected only for academic purposes. By doing this, the researcher protected the dignity of the students while effectively continuing the research process

### 3.5 Analytical Framework

This study is grounded in the Matrix Language Frame (MLF) model of code-switching. The Matrix Language Frame (MLF) model was proposed by Carol Myers-Scotton in her book "Duelling Languages: Grammatical Structure in Code-Switching (1993)" and later on, she refined her framework by adding the 4M model.

#### Matrix Language Frame (MLF) Model

Matrix Language Frame (MLF) model is a socio-linguistic theory which was developed to explain the intra-sentential code-switching, which means switching between two codes within the same sentence or clause. Mayers-Scotton argued that this code-switching is not random but structurally constrained and asymmetrical. By asymmetrical, she means grammatically constrained and predictable patterns of code-switching. Whenever the bilinguals switch codes, one language dominates the sentence by acting as the provider of grammatical structure, while another language acts as subordinate by playing the lexical role. This distinction leads to the emergence of two types of languages:

- **Matrix Language (ML):** According to Myers-Scotton, matrix language controls the grammatical structure of the sentence. It provides the morphosyntactic frame of the sentence and

determines the word order. It provides the system morphemes such as agreement markers, inflection affixes, agreement markers, tense markers, etc.

- **Embedded Language (EL):** Embedded language is the language inserted by the speakers while speaking. Instead of determining the grammatical structure of the bilinguals, it conforms to the grammatical structure provided by the matrix language. It provides the content morphemes such as noun, verb and adjective.

**Mayers-Scotton gave the two core principles:** The Morpheme Order Principle (MOP) and the System Morpheme Principle (SMP). The Morpheme order principle means if you mix the content word from the embedded language within the matrix language, then those content words will follow the rules of the matrix language. On the other hand, the system morpheme principle means that all the grammatical markers which determines the overall structure of the sentence, come from the matrix language.

### 4. Data Analysis

The study analyzed the code-switched data by using the Matrix Language Frame Model, which was developed by Myers-Scotton in the field of sociolinguistics. This framework has proposed two principles: the morpheme order principle and the system morpheme principle, which provide the in-depth exploration of the matrix and embedded language among the intermediates, undergraduates and graduates. For the analysis of interviews data, the thematic analysis approach is employed.

#### 4.1 Analysis of Intermediate Students' Code-Switched Clauses

The analysis of the data collected from the intermediate students showed consistent and systematic patterns across the dataset. Informed by the Matrix Language Frame Model, it indicates that Urdu operates as the matrix language while English functions as the embedded language within the clauses.

**Representative Example****Code-switched Clauses**

1. Main cricket khailna pasand kerta houn. (I love to play cricket)
2. Last week, main ny cricket tournament m participate b kiya tha. (I participated in the cricket tournament last week)
3. Hamaray coach ny humy kaha k cricket win kerny k liye practice aur focus boht zarori hai. (Our coach had asked that if we want to win the cricket, then we need to focus and to practice.)
4. Mein ub apni playing skills ko improve kerny k liye gym m strength training b start ker di hai. (Now I have joined a gym for strength training to improve my playing skills.)

Morpheme order of the clauses is subject-object-verb which shows that at the syntactic level, all the clauses produced by participant 1 adhere to the Urdu syntax. So, Urdu ultimately dominates the participant's English-Urdu code switching. The system morphemes such as 'main', 'ne', 'kerta houn', 'kiya tha', 'ny humy kaha', 'k liye', 'hai', 'ko', 'ker di hai', etc come from the Urdu language, which confirms the role of Urdu as the matrix language in the speaker's code-switched clause because the system morphemes determine the grammatical structure and regulate the placement of other elements in the clauses. On the other hand, English provides the content morphemes primarily, playing the role of embedded language. It provides nouns (cricket, cricket tournament, win, coach, gym, strength training) and verbs (participate, improve, start). These content morphemes provided by the English language are integrated into the Urdu grammatical elements without affecting its syntactic structure. This clearly aligns with the Matrix Language Frame Model, which argues that language that contributes system morphemes controls the grammatical structure of the code-switched clauses. Hence, by placing Urdu as the matrix language, the participant constructs his locally aligned identity rather than a modern and global one.

At the discourse function level, English seems to serve specific communicative needs. Multiple English content morphemes belong to the specific domain, such as sports (cricket, tournament), physical training (gym, strength training) and

action-oriented vocabulary (improve, participate). It indicates that code-switching serves the function of communicative efficiency and fills the lexical gap as no ready-made equivalent word is available for 'cricket', 'gym', 'strength-training', etc. These structural patterns are consistently observed across all the datasets of the intermediates' participants. They all shared the same feature where Urdu dominates in maintaining the grammatical structure while English remains at the level of embedded language and contributes only lexical items. This suggests a stable, recurring bilingual practice rather than random. All six participants tend to draw the vocabulary from different domains, but the overall structural organization of their speech remains the same: Urdu as the matrix language. Moreover, the linguistic patterns of the participants' code-switched clauses suggest that they actively construct their locally oriented identity through Urdu as the matrix language choice. This is evidently reflected in their reliance on Urdu grammatical structure. On the other hand, integration of English lexical items in the Urdu indicates the emerging engagement with global or academic domains. Hence, Urdu becomes the linguistic ground which is maintained through the insertion of selective English lexical items.

**4.2 Analysis of the Under-Graduate Students Code-Switching**

The analysis of undergraduates' code-switched clauses indicates more variable and transitional patterns as compared to the intermediate level. Guided by the Matrix Language Frame Model, the matrix language choice is not consistent across the data and English and Urdu both function alternatively as the matrix language within the code-switched clauses of the same speaker.

**Representative Example**

5. I think, charity shouldn't be considered as sirf religious commitment, in fact it's the zati responsibility of every person. (sirf = just, zati = personal)
6. Jub log zakat donate kerty hain to onka focus sirf short term help hota hai. (When people donate zakat, their focus is just short-term help.)

7. Government should provide free educational access to the zarorat-mund people because education empowers the people by increasing the employment opportunities. (zarorat mund =needy)

In the utterances of the speaker, there is a kind of inconsistency in the choice of matrix language. In two clauses, the matrix language is English, and in one clause the matrix language is Urdu. Morpheme order of the clause 1 and 3 is subject-verb-object, and all the system morphemes stem from the English language. In clauses 1 and 3, Urdu as the embedded language provides the content morphemes such as 'zati', 'sirf' and 'zarorat mund' while English controls the sentence structure by providing the functional morphemes such as auxiliary (should, shouldn't), articles (the), subject verb agreement (education empowers, etc.), conjunction (because) and preposition etc. On the other hand, clause 2 has Urdu as the dominant matrix language, which controls the syntactic structure and grammar and makes the content morphemes, which stem from English, structurally constrained. In clause 2, the content morphemes are 'donate', 'focus' and 'short-term help' while system morphemes are 'jub', 'hain', 'hota hai', etc. But the consistency in the accurate syntactic structures either SVO or SOV, shows that code-switching is not random, while inconsistency in the choice of this participant seems to indicate the influence of the level of education. Similar patterns were observed across the participant 1, 2 and 4 which suggests that this variability is a characteristic of the undergraduate group.

Data where English dominates as the matrix language indicates that at this level, the discourse appears to function more formally and abstractly. On the other hand, the Urdu-dominant clauses are used for more elaborate and explanatory segments. Participants 1, 2 and 4 seem to reflect a more hybrid or transitional identity orientation as the speakers don't align with a single linguistic identity and navigate between Urdu and English. They use both languages for structural and lexical resources. It suggests a more evolving identity where local and global linguistic identity co-exist. Along with those participants' whose English and Urdu both appeared as the matrix language, there are the

students for whom only Urdu consistently appeared as the matrix language.

#### Representative Example

8. University life challenging b hai aur exciting b hai. (University life is challenging and exciting too).

9. Pora time-period assignments, quizzes, presentation aur makeup classes hi busy rakhti hai lekin friends aur campus activities excitement b create kerti hain. (Whole time-period assignments, quizzes, presentation and make-up classes keep us busy but friends and campus activities create the excitement as well).

Both clauses have the subject-object-verb morpheme order and system morphemes stems from the Urdu language, making it the matrix language, while English is the embedded language, which only provides the content morpheme and is bound to follow the morpheme order of the matrix language. In these clauses, the system morphemes are 'hai', 'aur', 'hai lekin', 'kerti hain', 'b' etc., and they determine the grammatical structure of the clauses while the content morphemes are adjective (challenging, exciting, busy), noun (assignments, quizzes, presentation, campus activities, university life) and verb (create). Code-switching is structurally constrained. Similarly, the consistent dominance of Urdu as the matrix language is observed in participants 5 and 6. In the representative example, the discourse functions of the English content morphemes seem to be more associated with academic settings where English terms are more easily and readily available. However, the structure of the clause is still controlled by the Urdu language. Moreover, the dominance of Urdu suggests that speakers tend to align their linguistic identity with the local language while using English to fill the lexical gaps. Thus, at the undergraduate level, some participants demonstrate English as the matrix language while some demonstrates Urdu. This variability of the matrix language choice across the six participants indicates that matrix language is not uniform at this level, but constantly evolving, which ultimately leads to the identity variation as well.

### 4.3 Analysis of the Graduate Student Code-Switching

The analysis of the English-Urdu code-switched clauses of the graduates' students shows highly consistent pattern across the whole data. Guided by the Matrix Language Frame Model, English remains the matrix language of all the six participants while Urdu serves as the embedded language.

#### Representative Example

10. In Pakistan, climate is rapidly changing due to the lack of shoar among the people because anjanay main they throw the wrappers and dirty waste where they want which can lead to the shadeed climate issues. (shaour = sense/awareness, anjanay main = unknowingly, shadeed = severe).

11. It can lead to unpredictable mosam and tofani barishain which ultimately lead to the seelab and other natural disasters. (mosam = season, tofani barishain = torrential rains, seelab = flood).

12. Not only Pakistan, the whole world will face the serious khatra if they don't overcome this issue as soon as possible. (khatra = danger)

All three clauses contain English as the matrix language because the morpheme order of the clauses is subject-verb-object, and the system morphemes such as auxiliary (is, can, will), preposition (in, due to, among), tense marker, conjunction (because, if) and subject-verb agreement (climate is, world will face) stem from the English language. Urdu here functions as the embedded language by providing the content morphemes, such as 'shaour', 'shadeed', 'anjanay main', 'mosam', 'tofani barishain', 'seelab' and 'khatra' while English determines the grammatical structure as well as the position of the content morphemes. The similar morpheme order is observed in the code-switched clauses of participant 2,3, 4, 5 and 6. From a discourse-functional perspective, English is the matrix language used to discuss more societal and global issues such as climate change. Urdu lexical items appear to be inserted for emphasis and specificity. These patterns are consistently observed in the entire graduate participants' data, which indicates a stable preference to English, unlike the undergraduate participants, where there was

variability in the choice of the matrix language. Hence, these consistent patterns suggest an inclination towards a more English-dominant, global and academically grounded identity. At the same time, the insertion of Urdu lexical items indicates the configuration of a hybrid identity with a stronger academic and global orientation.

### 4.4 Analysis of Semi- Structured Interview Data

#### 4.4.1 Theme: Educational Exposure and the Matrix Language Choice

The semi-structured interviews conducted with the participants of intermediate, undergraduates and graduates indicate that education plays a very significant role in shaping the code-switching patterns, particularly the choice of matrix language. Most of the participants reported that their language use evolved over time due to the evolving educational exposure.

**Participant 2 (intermediate):** "Muje English grammar ka achy sy knowledge hai lekin abi itna educational emphasize nai hai jiski waja sy naturally Urdu ziyada comfortable lagti"

**Participant 1 (under-graduate):** "I didn't exhibit the same language use at the intermediate level, but university education demands more English proficiency as compared to Intermediate. So, over the time its gradually evolving."

**Participant 6 (Graduates):** "Obviously we are at the graduate level and have gone through and is going through English-dominant education and research articles also demand English. So, now at this level we feel naturally oriented towards English while still use Urdu but it's more often to fill the lexical gap or more cultural emphasize."

The responses of all the participants at each level suggest educational influence on the choice of matrix language in their communication. The data indicates that undergraduates' participants did not have the same tendency to matrix language at the intermediate level, and graduates did not have at the undergraduate level, which clearly highlights the influence of education on the choice of matrix language. University education in Pakistan demands more English as compared to college and school education. The data suggests that increased engagement with the English language effects the matrix language choice.

## 5. Findings and Discussion

The analysis of the English-Urdu code-switching across three educational levels: intermediates, undergraduates and graduates, reveals a very systemic shift in the matrix language among the bilinguals and the code-switching is structurally constrained. Among the intermediates, Urdu consistently dominates as the matrix language despite the competence of both languages. Among the undergraduates, English began to emerge as the matrix language in the participants' code-switched clauses, but this happens very inconsistently because in some clauses, Urdu dominates as the matrix language. On the other hand, at the graduate level, English participants consistently demonstrate English as the matrix language in their code-switching, while Urdu merely serves the content morphemes. From the perspective of identity construction, there is a parallel shift in the identity orientation across the educational level. At the intermediate level, participants exhibit more locally grounded linguistic identity. On the other hand, variability in the undergraduates' matrix language choice indicates a transitional identity where speakers negotiate between locally and globally associated linguistic identity. In contrast, at the graduation level, consistent use of English as the matrix language suggests an increased alignment with the academically oriented and English-dominant linguistic identity. However, they still selectively retained lexical items from the Urdu language. Moreover, the interview data showed that education remains a significant factor in the choice of matrix language. Increased exposure to English-dominant education, increases the tendency towards the dominance of English as the matrix language while Urdu still plays a lexical role.

The study confirmed the Myer-Scotton Matrix Language Frame Model (1993), which emphasizes that code-switching is structurally constrained. All the code-switched clauses among the intermediates, undergraduates and graduates were grammatically structured rather than arbitrary. While there was systematic variation in the choice of matrix language among the intermediates, undergraduates and graduates, because at each level there is a slight variation in the language that controls the

grammatical structure. All the intermediate students consistently used Urdu as the matrix language despite the sufficient competence in both languages, as Yang, Rauwolf, Frances, Wei, Molina-Nieto, Dunabeitia and Thierry (2023) revealed in their study on Chinese-English bilinguals that bilinguals frequently preferred Chinese over English despite the competence in both languages. Similarly, Puig-Mayenco, Cunnings, Bayram, Miller, Tubau and Rothman (2018) found the dominance of L1 among the bilinguals who had the competence of both Spanish, which supports the present finding, that despite the competence in English and Urdu, intermediate students consistently tend towards projecting their local linguistic identity. Despite the preference for L1, they switch to English for content morphemes, as Catalan. Nawaz, Yousaf and Jabeen (2023) identified that English is frequently present in the ESL classrooms of Pakistan. On the other hand, among the undergraduates English began to appear as the matrix language due to the influence of social context while among the graduates English consistently appeared as the matrix language due to social factors, as Mann and Bruin (2021) examined that bilinguals vary in their language preferences depending on the context and experience. Similarly, Fasya and Sari (2021) support the argument that socio-cultural factors influence the language preference among multilinguals. Similarly, Riaz (2026) investigating the role of English-Urdu code-switching and highlighted its role in shaping the identities.

The main driver of the English as the matrix language over the period of time is education because educational institutions emphasize English over any other languages, as Ali (2026) argued that Pakistani education highly values English and marginalizes the regional languages. Moreover, Ntombela (2023) also argued about English hegemony in the higher educational system. So, the student step into higher education, the influence of English hegemony increases which leads to dominance of English in code-switching. In the present study, students over the period of time preferred English as the matrix language because English carries more prestige as the educational exposure increases. Raza, Imran, Abid

and Nadeem (2025) have long argued in their study that English carries more prestige and value as compared to other languages, creating a diglossia in the society.

The findings of the present study demonstrate a clear inter-relationship between identity construction, matrix language and education. As the matrix language choice across the three educational level vary, which indicates the influence of education. This observation is further supported by the semi-structured interview, which suggests that education directly influences matrix language choice. This ultimately influences identity construction, as a shift in educational level leads to a shift in matrix language, which in turn shifts the identity construction. Bucholtz and Hall (2005) emphasized that identity negotiation is a linguistic phenomenon and it is shaped by the social context. Moreover, Usmanova (2024) also examined how sociolinguistic factors influence the social identities. Furthermore, this study has a theoretical implication because it shows the successful application of the Matrix Language Frame Model on code-switching and confirms that code-switching is not random but structurally constrained. Moreover, its two principles are applicable to English-Urdu code-switching analysis.

## 6. Conclusion

The present study aimed to investigate the matrix language among the bilinguals at the three educational levels: intermediates, undergraduates and graduates. Moreover, the study extends to investigate the relationship between identity construction, matrix language and education. Through the analysis of the collected data from the participants, the study revealed the systematic variations in the choice of the matrix language due to the impact of education. Intermediate students performed more locally grounded linguistic identity by preferring Urdu as the matrix language, and it occurred due to the limited educational and institutional emphasize, under-graduate students inconsistently chose English as the matrix language and projected more transitional linguistic identity, while graduates consistently prefer English as the matrix language, while Urdu only works as the embedded language, which showed more stable

academically grounded linguistic identity. Moreover, the findings clearly demonstrated a correlation between identity construction, matrix language and education. Education appeared to be the most independent factor that influences the matrix language while ultimately influencing the identity negotiations as language is the source through which identity is expressed and negotiated.

## REFERENCES

- Ahmad, F., Muhammad, A. M., & Shahid, A. (2023). Code-switching and code-mixing among bilinguals: A socio-pragmatic study. *Pakistan Social Sciences Review*, 7(3), 35-48. [https://doi.org/10.35484/pssr.2023\(7-III\)04](https://doi.org/10.35484/pssr.2023(7-III)04)
- Ali, R. (2026). 'We are still British or American colony, we are still a colony, we are not free': language ideologies, policies in education in Pakistan. *Current Issues in Language Planning*, 27(1), 67-90. <https://doi.org/10.1080/14664208.2024.2447972>
- Ambarita, E., & Meliala, G. C. A. S. (2024). Types of code-switching in Kompas newspaper: A study in political and law articles. *Journal of Applied Linguistics*, 4(1), 12-23. <https://doi.org/10.52622/joal.v4i1.236>
- Asgher, T., Shahzad, A. K., & Hanif, A. (2023). Unveiling Pakistani English: A Linguistic and Cultural Odyssey of Features, Evolution and Distinctiveness. *Qlantic Journal of Social Sciences and Humanities*, 4(3), 299-311.
- Beardmore. (1982). *Bilingualism. Basic principles*. Clevedon: Multilingual Matters
- Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociolinguistic approach. *Discourse Studies*, 7(4-5), 585-614
- Chaer, A. (1995). *Sosiolinguistik: Suatu Pengantar*. Jakarta: PT. Rineka Cipta.
- Cormack, Mc W.C. and Stephena A.W. (1979). *Language and Society*. New York: Mountain Publisher.
- Crystal, D. (1992). An encyclopedic dictionary of language and languages. (No Title).

- De Fina, A. (2007). *Code-switching and the construction of ethnic identity in a community of social practice*. *Language in Society*, 36(3), 371-392
- Fasya, M., & Sari, D. G. (2021). *Sociocultural factors that determine language choice in a multilingual society*. In *Proceedings of the Fifth International Conference on Language, Literature, Culture, and Education (ICOLLITE 2021)* (Vol. 595, pp. 412-418). Atlantis Press. <https://doi.org/10.2991/assehr.k.211119.064>
- Fatima, M., & Amin, S. T. U. (2025). Student's Perception of Teachers' Code-Switching in Pakistani English Language Classroom. *International Journal of Social Sciences Bulletin*, 3(6), 848-855. <https://ijssbulletin.com/index.php/IJSSB/article/view/948>
- Graddol, D., Leith, D., & Swann, J. (Eds.). (1996). *English: history, diversity, and change* (Vol. 1). Psychology Press.
- Grosjean, F. (2014). Processing mixed language: Issues, findings, and models. In *Tutorials in Bilingualism* (pp. 225-254). Psychology Press.
- Gumperz, J. (1977). *Conversational Code-Switching in Discourse Strategies*. New York: Holt, Rinehart and Winston
- Gumperz, J. J. (1982). *Discourse strategies* (No. 1). Cambridge university press
- Hadei, M., Kumar, V. C., & Koik, S. J. (2016). Social factors for code-switching – A study of Malaysian-English bilingual speakers. *International Journal of Language and Linguistics*, 4(3), 122-127. <https://doi.org/10.11648/j.ijll.20160403.15>
- Hammink, J. E. (2000). A comparison of the code-switching behaviour and knowledge of adults and children. Retrieved October 20, 2010 from [http://hamminkj.tripod.com/babel/CS\\_paper.htm](http://hamminkj.tripod.com/babel/CS_paper.htm).
- Heller, M. (2007). Bilingualism as ideology and practice. In *Bilingualism: A social approach* (pp. 1-22). London: Palgrave Macmillan UK.
- Holmes, J. (2001). *An Introduction to Sociolinguistics*. London: Longman.
- Huttener, A. (1997). *Code-switching and the Classroom Teacher*. *Language in Education*.
- Iqra, U., Mahmood, A., Firdaus, A., & Saqlain, F. (2025). Code-switching and code-mixing: A sociolinguistics study of Nimra Ahmed's Maala. *International Journal of Social Sciences Bulletin*, 3(10), 544-551. <https://ijssbulletin.com/index.php/IJSSB/article/view/1376>
- Kachru, B. B. (1986). *The alchemy of English: The spread, functions and models of non-native Englishes*. Oxford: Pargamon Press.
- Khan, A. A., Nazir, N., Hamid, A., & Nazir, N. (2021). Social factors and Pashto-English code-switching: The case of English as embedded language in Khyber Pakhtunkhwa, Pakistan. *Humanities & Social Sciences Reviews*, 9(3), 1673-1681. <https://doi.org/10.18510/hssr.2021.93169>
- Mann, A., & de Bruin, A. (2022). Bilingual language use is context dependent: using the Language and Social Background Questionnaire to assess language experiences and test-retest reliability. *International Journal of Bilingual Education and Bilingualism*, 25(8), 2886-2901. <https://doi.org/10.1080/13670050.2021.1988049>
- Myers-Scotton, C. (1993). *Duelling Languages: Grammatical Structure in Codeswitching*. Clarendon Press.
- Myers-Scotton, C. (1997). *Duelling languages: Grammatical structure in codeswitching*. Oxford University Press.
- Namba, K. (2004). An overview of Myers-Scotton's matrix language frame model. *Senri International School (SIS) Educational Research Bulletin*, 9, 1-10. Retrieved from <http://yayoi.senri.ed.jp/research/re09/namba.pdf>
- Nilep, C. (2006). "Code switching" in sociocultural linguistics. *Colorado research in linguistics*. <https://philpapers.org/archive/nilcsi.pdf>

- Ntombela, B. X. S. (2023). The sociolinguistic problems of English medium instruction in the Middle East and North Africa: Implications for epistemic access. *Frontiers in Psychology*, 14, 1084626.
- Poplack, S. (1980). Sometimes I'll Start a Sentence in Spanish Y Termino en español: Toward a Typology of Code-Switching. *Linguistics*, 18, 581-618.  
<http://dx.doi.org/10.1515/ling.1980.18.7-8.581>
- Poplack, S. (2020). Sometimes I'll start a sentence in Spanish y termino en español: toward a typology of code-switching. In *The bilingualism reader* (pp. 213-243). Routledge.
- Pride, J.B. and Holmes J. (1979). *Sociolinguistics*. Great Britain
- Prosper, A. (2025). A descriptive analysis of sociocontextual factors influencing code switching behaviour among university students. *Canadian Journal of Educational and Social Studies*, 5(2), 1-20.  
<https://doi.org/10.53103/cjess.v5i2.333>
- Puig-Mayenco, E., Cunnings, I., Bayram, F., Miller, D., Tubau, S., & Rothman, J. (2018). Language dominance affects bilingual performance and processing outcomes in adulthood. *Frontiers in Psychology*, 9, 1199.  
<https://doi.org/10.3389/fpsyg.2018.01199>
- Rahman, T. (1990). *Pakistani English: The linguistic description of a non-native variety of English* (Vol. 3). National Institute of Pakistan Studies, Quaid-i-Azam University.
- Raza, A., Imran, A., Abid, M. A., & Nadeem, M. (2025). The Prestige of English vs. Urdu: A Sociolinguistic Analysis of Diglossia and Social Class in Pakistan. *Journal of Applied Linguistics and TESOL (JALT)*, 8(2), 2702-2715.
- Riaz, B. (2026). Identity construction through Urdu-English code-switching in mobile network advertisements in Pakistan. *Liberal Journal of Language & Literature Review*.  
<https://www.researchgate.net/publication/400310940>
- Tabbasum, R. (2023). Quantitative investigation of students' perceptions regarding code switching in Pakistani English classrooms at university level. *Journal of English Language, Literature and Linguistics (JELLL)*, 1(1), 53-72.  
Retrieved from <https://ojs.aiou.edu.pk/index.php/jelll/article/view/1438>
- Tahir, A., Fatima, I., & Abuzar, N. (2016). Teachers' and students' attitude toward code alternation in Pakistani English classrooms. *JEELS (Journal of English Education and Linguistics Studies)*, 3(1), 85-108.
- Usmanova, N. X. (2024). Sociolinguistic factors in code switching. *American Journal of Interdisciplinary Research and Development*, 29, 220-224.  
<https://ajird.journalspark.org/index.php/ajird/article/view/1229/1181>
- Valdes-Fallis, G. (1978). Code Switching and the Classroom Teacher. *Language in Education: Theory and Practice*, No. 4.
- Wei, L. (Ed.). (2000). *The bilingualism reader* (Vol. 11). London: Routledge.
- Yang, W., Rauwolf, P., Frances, C., Wei, Y., Molina-Nieto, O., Duñabeitia, J. A., & Thierry, G. (2023). Evidence for Strategic Language Use in Chinese-English Bilinguals. Available at SSRN 4396191.
- Zainab, A., Ahmed, I., & Kashif, W. (2024). Urdu-English code-switching in Pakistani English. *Voyage Journal of Educational Studies (VJES)*, 4(1), 47-56.  
<https://share.google/DbZ9x4hVoDToxs250>